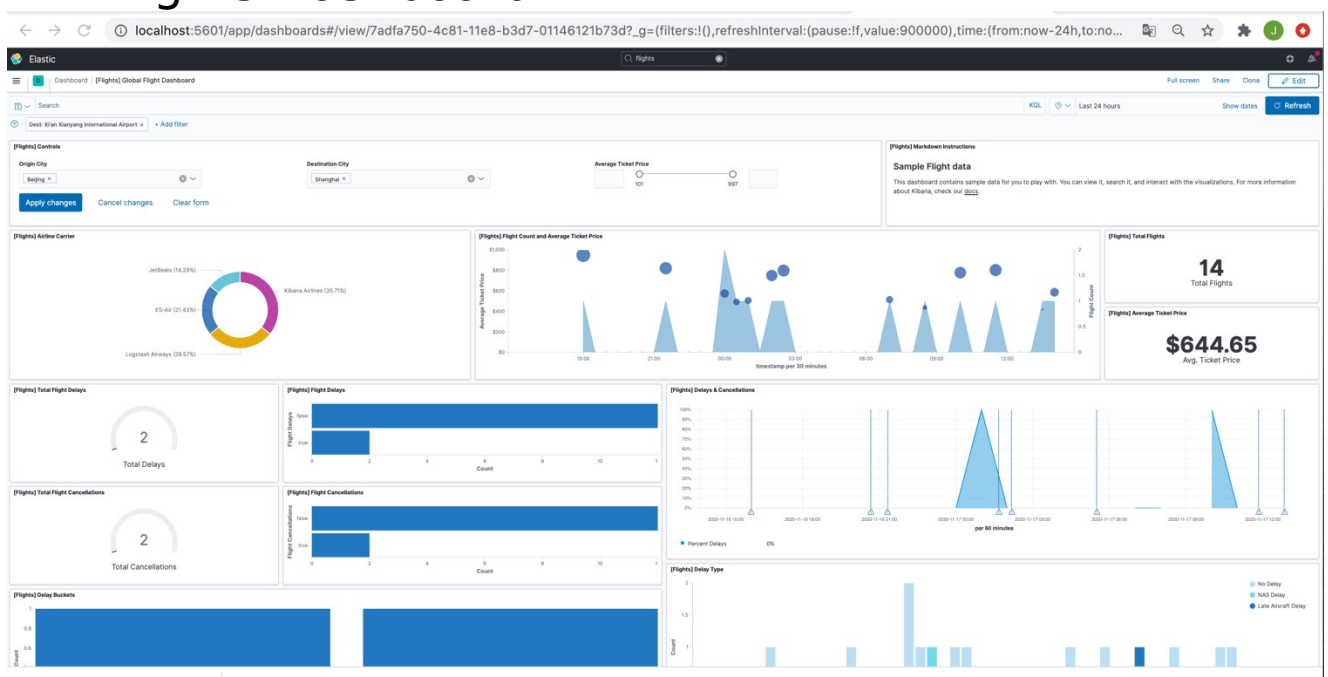Nicolas Boussenina
Nicolas Oulianov
Julien Liu

Real time Big Data Search & Analytics
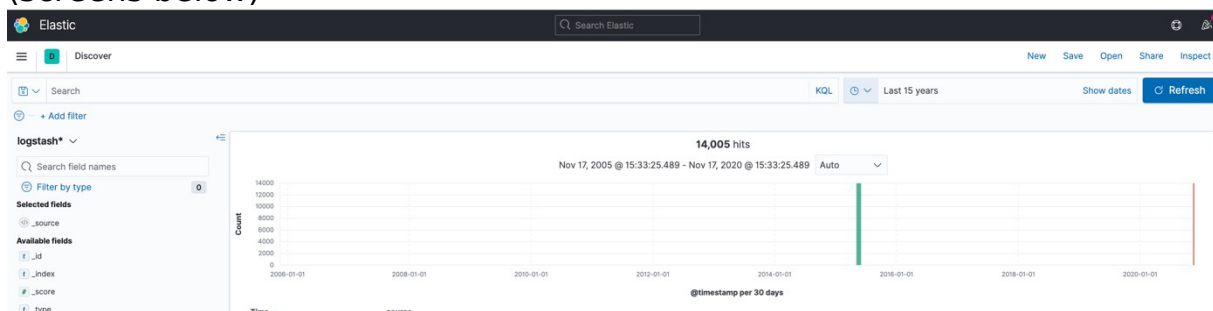
Lab 1

1.1 Flights Dashboard

1.1 Please specify in your report the total number of documents of each index.

ba* : 1 000hits
logstash* : 14 005 hits
shakes* : 111 396 hits
(screens below)

1.2 Please do some screenshots of at least 1 search or filter you applied on each index.

1.3 Do a screenshot of the dashboard and add it to your report:

+ Add filter

Bar Example

Speaking Parts

Play Name

Markdown Example

This is a tutorial dashboard!

Lab 1 : M2 Data Science IPParis + Nicolas B, Nicolas O, Julien

Max Example

LAYERS

logstash*

Road map

Pie Example

≥ 1,000 and < 2,9...
≥ 3,000 and < 6,...
≥ 7,000 and < 14,...
≥ 15,000 and < 3...
≥ 31,000 and < 5...
35
22

Lab 2

2.1 Do a screenshot of the monitoring and add it to your report:
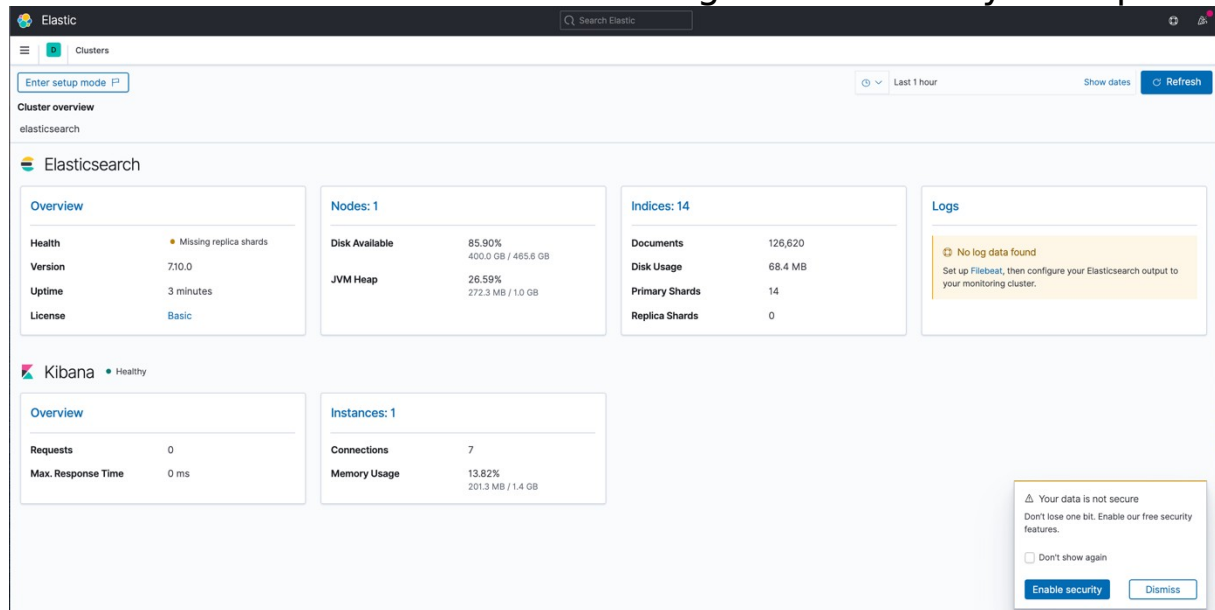
What is their status? Do they have unassigned shards?

Their status is "Yellow" and they have unassigned shards.

2.2 Do a screenshot of indices monitoring and add it to your report

Why is there unassigned shards for all our indices created in Lab 1 ?

There are unassigned shards because we only have 1 node, so replica shards aren't assigned to any nodes.

2.3 Do a screenshot of the monitoring of the "bank" index and add it to your report
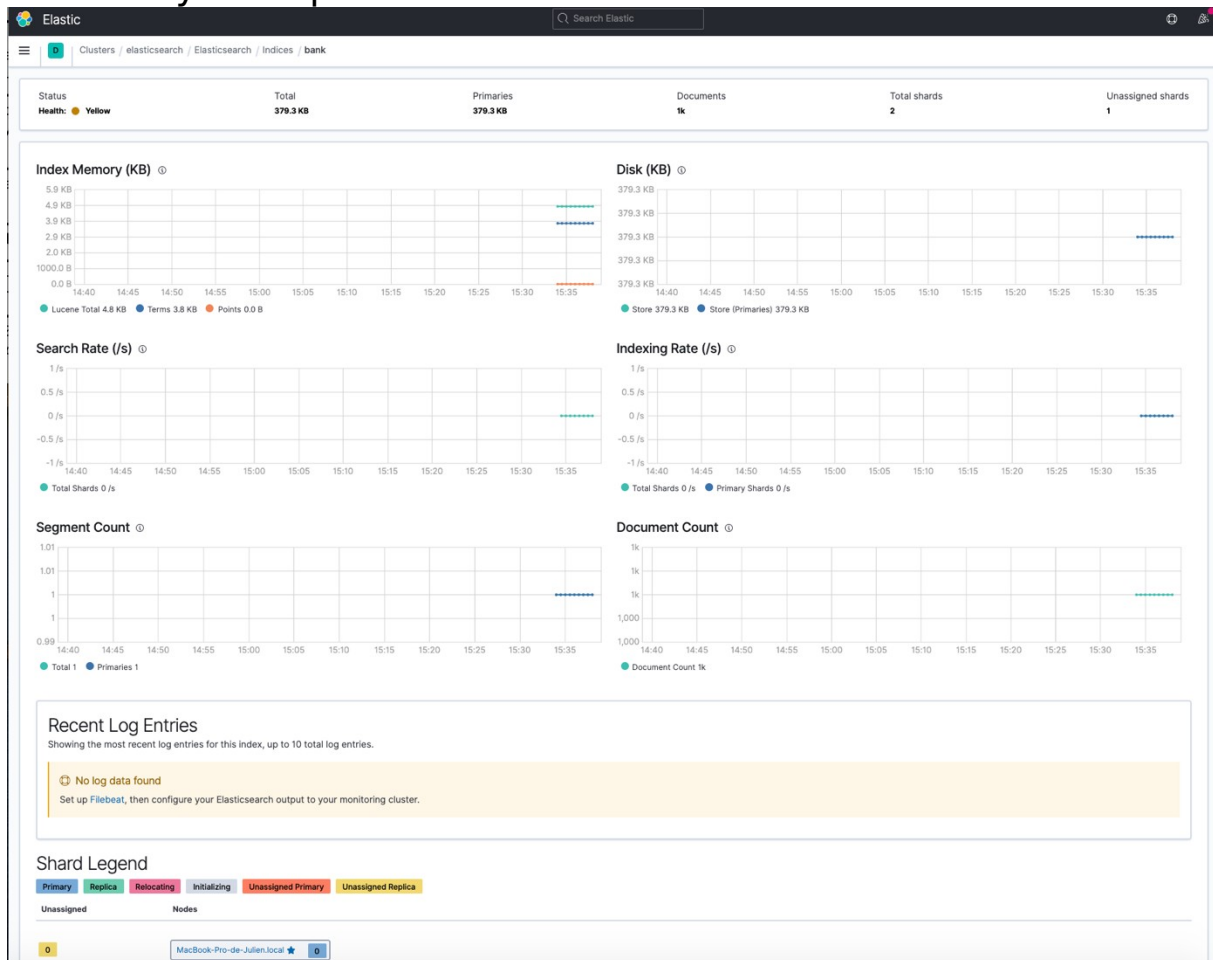
2.4 Do a screenshot of the monitoring of the "bank" index, after the addition of 1 node and add it to your report. Comment about what happen

on the bank index monitoring page ?
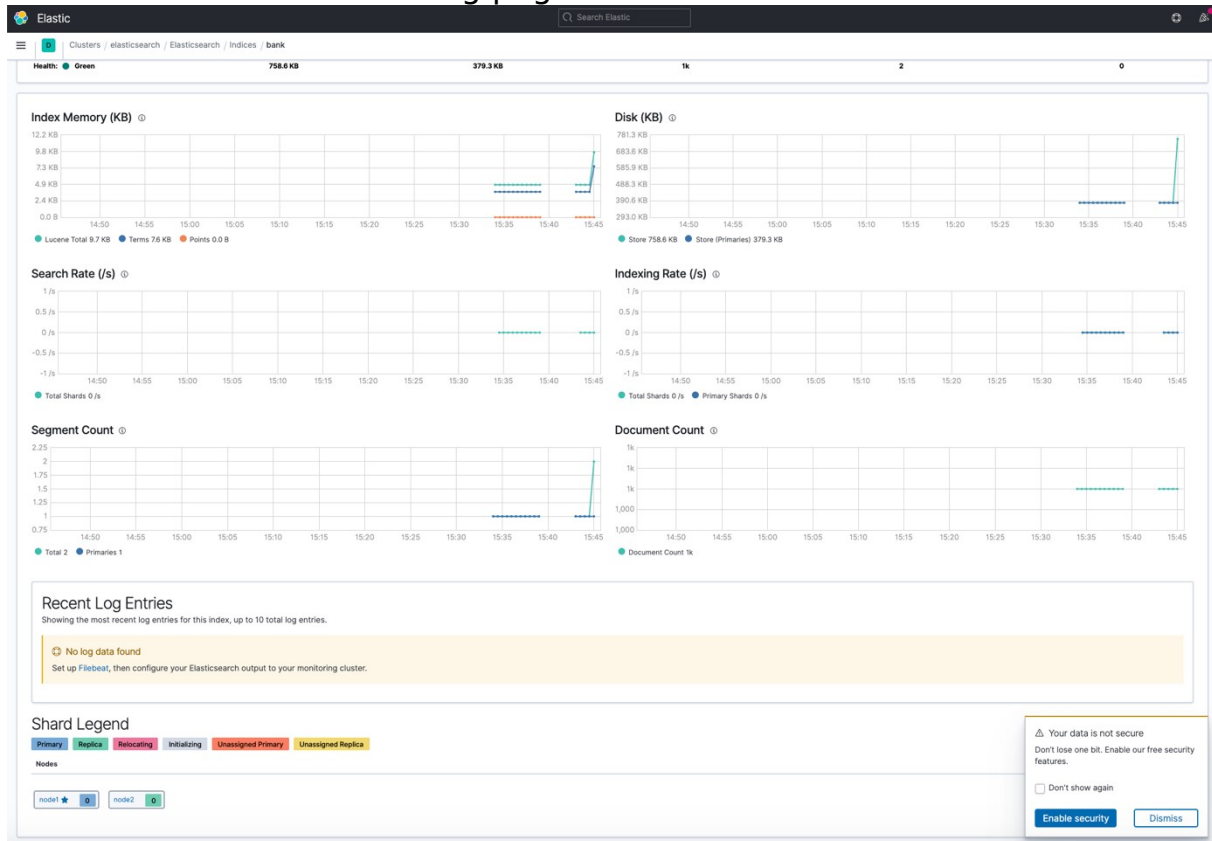
We added a node to our cluster. Now, all the shards are assigned to a node. We can see on the screenshot that the primary shard is assigned to node1 and the replica to node2.

2.5 Do a screenshot of the monitoring of the "bank" index, after the addition of 2 nodes and add it to your report. Comment about what happen on the bank index monitoring page?
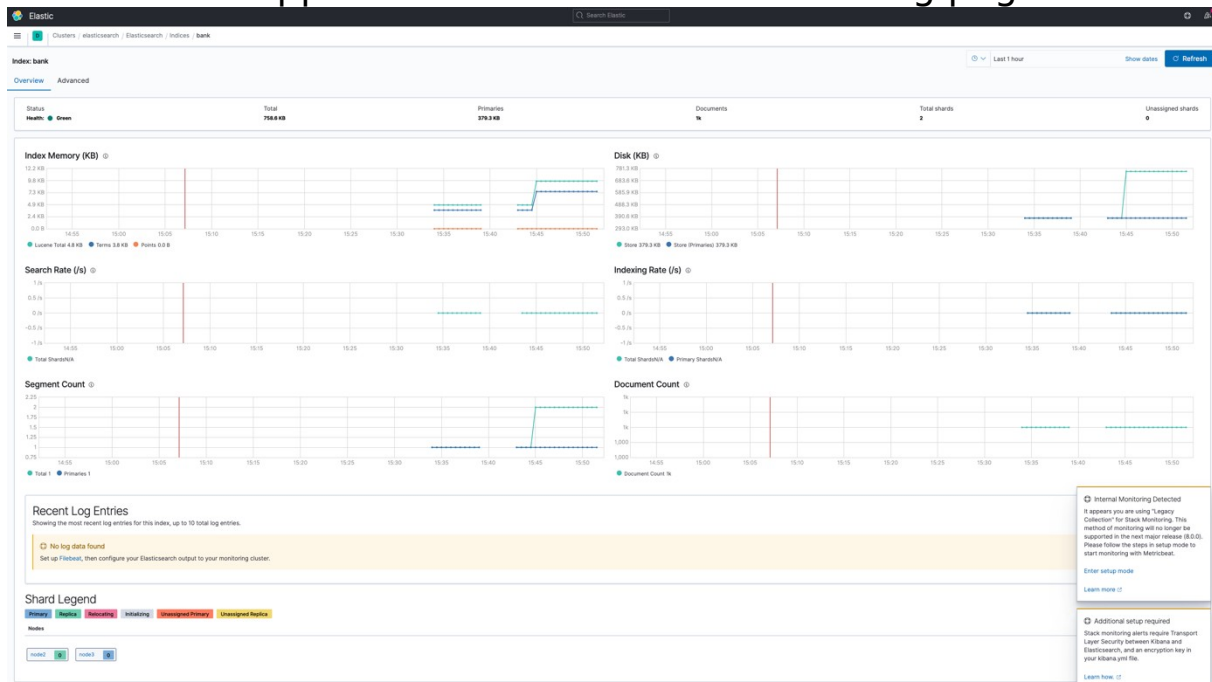
The number of replicas for the index "bank" is set to 1. Here, the primary shards is assigned to node3 and the replica shard to node2. We do not see node1 here because we only have 1 replica.

2.6 Do a screenshot of indices monitoring, after the addition of 2 nodes and add it to your report
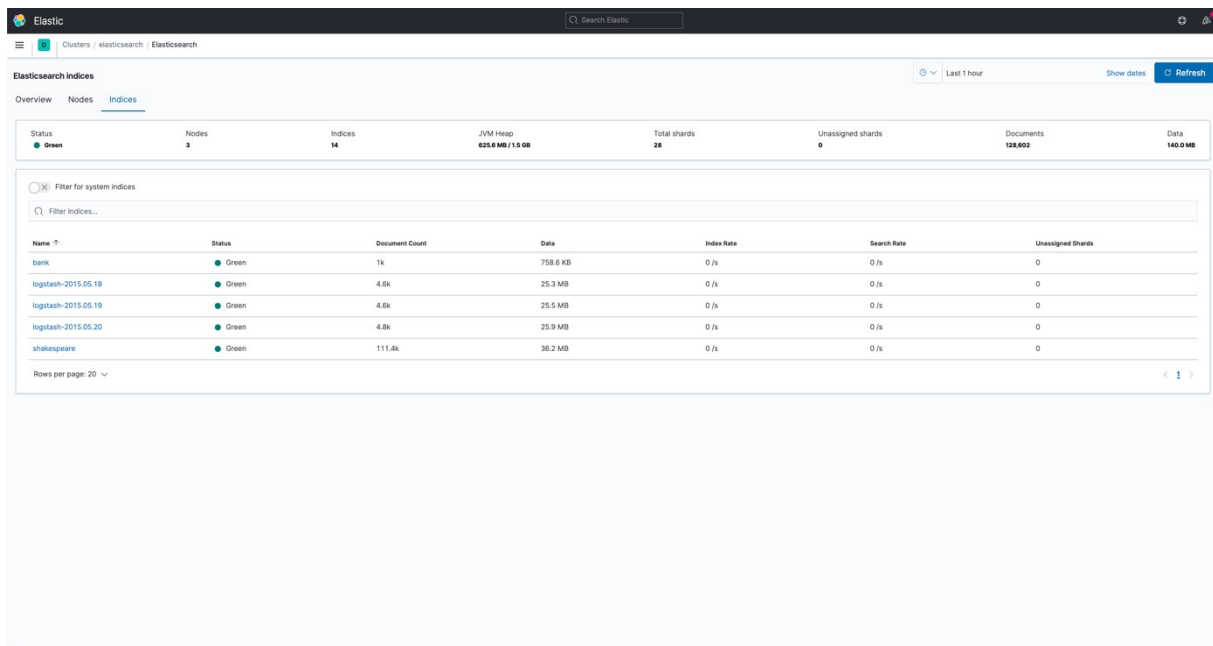
2.7 Do a screenshot of the monitoring of the "bank" index, after reducing number of replica to 0 and add it to your report. What happened ?
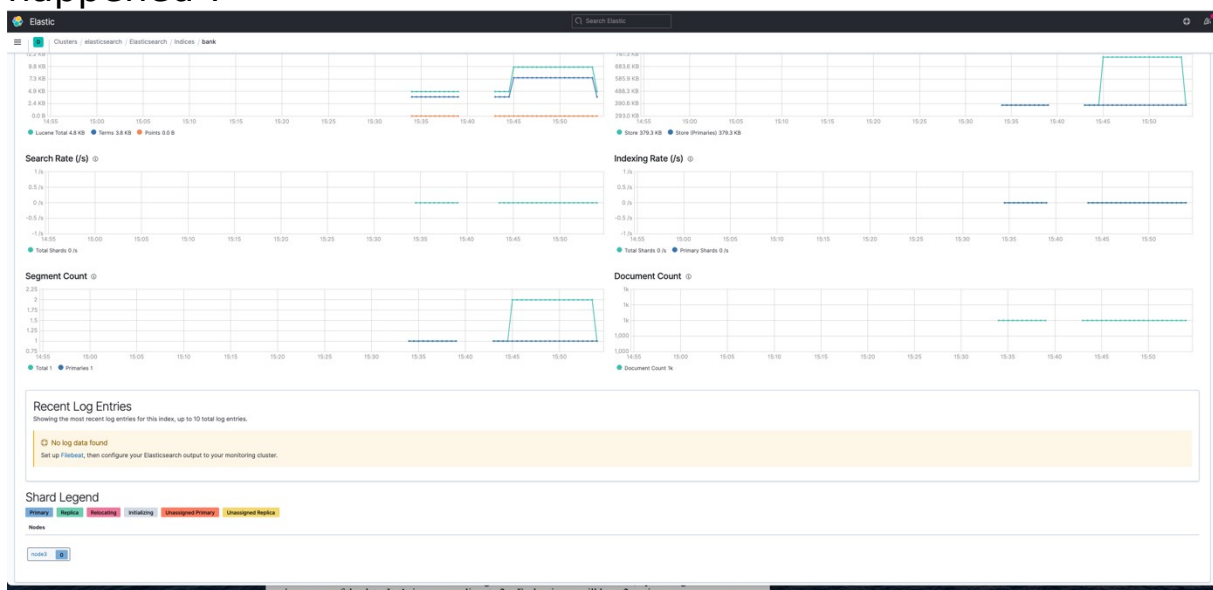
We set the number of replica to 0. So, we only have 1 primary shard assigned to node3 for bank index.

2.8 Do a screenshot of the monitoring of the "bank" index, after increasing number of replica to 2 and add it to your report. What happened ?
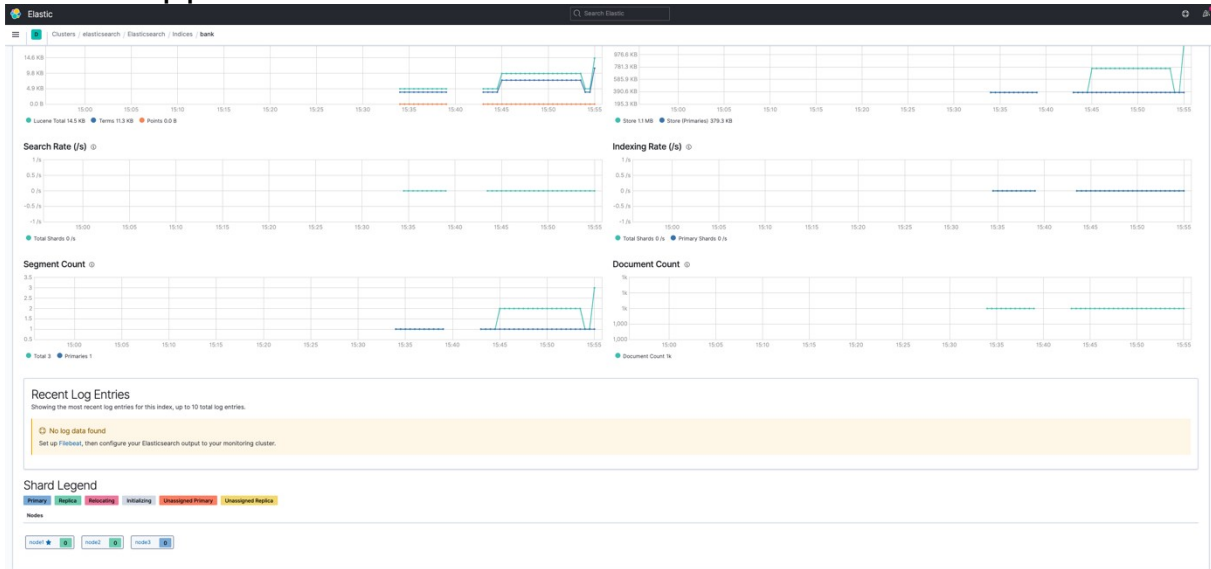
We have to 2 replicas. Now the primary shard is assigned to node3 and one replica is assigned to node2 and another replica to node1.

Lab3

3.1 Do a screenshot of your search results and add it to your report

```
1 ▾ {
2      "took" : 40,
3      "timed_out" : false,
4 ▾    "_shards" : {
5        "total" : 2,
6        "successful" : 2,
7        "skipped" : 0,
8        "failed" : 0
9 ▴    },
10 ▾   "hits" : {
11 ▾     "total" : {
12          "value" : 10000,
13          "relation" : "gte"
14 ▴     },
15        "max_score" : 16.079674,
16 ▾     "hits" : [
17 ▾       {
18            "_index" : "shakespeare-catalog",
19            "_type" : "_doc",
20            "_id" : "19",
21            "_score" : 16.079674,
22 ▾         "_source" : {
23              "type" : "line",
24              "line_id" : 20,
25              "play_name" : "Henry IV",
26              "speech_number" : 1,
27              "line_number" : "1.1.17",
28              "speaker" : "KING HENRY IV",
29              "text_entry" : "The edge of war, like an ill-sheathed knife,"
30 ▴         }
31 ▴       },
32 ▾       {
33            "_index" : "shakespeare-catalog",
```

```
34            "_type" : "_doc",
35            "_id" : "14215",
36            "_score" : 12.424048,
37 ▾         "_source" : {
38              "type" : "line",
39              "line_id" : 14216,
40              "play_name" : "Alls well that ends well",
41              "speech_number" : 2,
42              "line_number" : "3.3.7",
43              "speaker" : "BERTRAM",
44              "text_entry" : "To the extreme edge of hazard."
45 ▴         }
46 ▴       },
47 ▾       {
48            "_index" : "shakespeare-catalog",
49            "_type" : "_doc",
50            "_id" : "93126",
51            "_score" : 12.424048,
52 ▾         "_source" : {
53              "type" : "line",
54              "line_id" : 93127,
55              "play_name" : "The Tempest",
56              "speech_number" : 4,
57              "line_number" : "4.1.31",
58              "speaker" : "FERDINAND",
59              "text_entry" : "The edge of that days celebration"
60 ▴         }
61 ▴       },
62 ▾       {
63            "_index" : "shakespeare-catalog",
64            "_type" : "_doc",
65            "_id" : "32995",
66            "_score" : 11.832202,
```

Add the exact number of hits in your report that you will get from this request :

The exact number of hits from this request is 32484 (screen below)

```
1 ▾ {
2     "took" : 7,
3     "timed_out" : false,
4 ▾   "_shards" : {
5       "total" : 2,
6       "successful" : 2,
7       "skipped" : 0,
8       "failed" : 0
9 ▴   },
10 ▾  "hits" : {
11 ▾    "total" : {
12         "value" : 32484,
13         "relation" : "eq"
14 ▴    },
15      "max_score" : 16.079674,
16 ▾    "hits" : [
17 ▾      {
18           "_index" : "shakespeare-catalog",
19           "_type" : "_doc",
20           "_id" : "19",
21           "_score" : 16.079674,
22 ▾         "_source" : {
23             "type" : "line",
24             "line_id" : 20,
25             "play_name" : "Henry IV",
26             "speech_number" : 1,
27             "line_number" : "1.1.17",
28             "speaker" : "KING HENRY IV",
29             "text_entry" : "The edge of war, like an ill-sheathed knife,"
30 ▴         }
31 ▴       },
32 ▾       {
33           "_index" : "shakespeare-catalog",
34           "_type" : "_doc",
```

3.2 Do a screenshot of your search profiler and add it to your report :
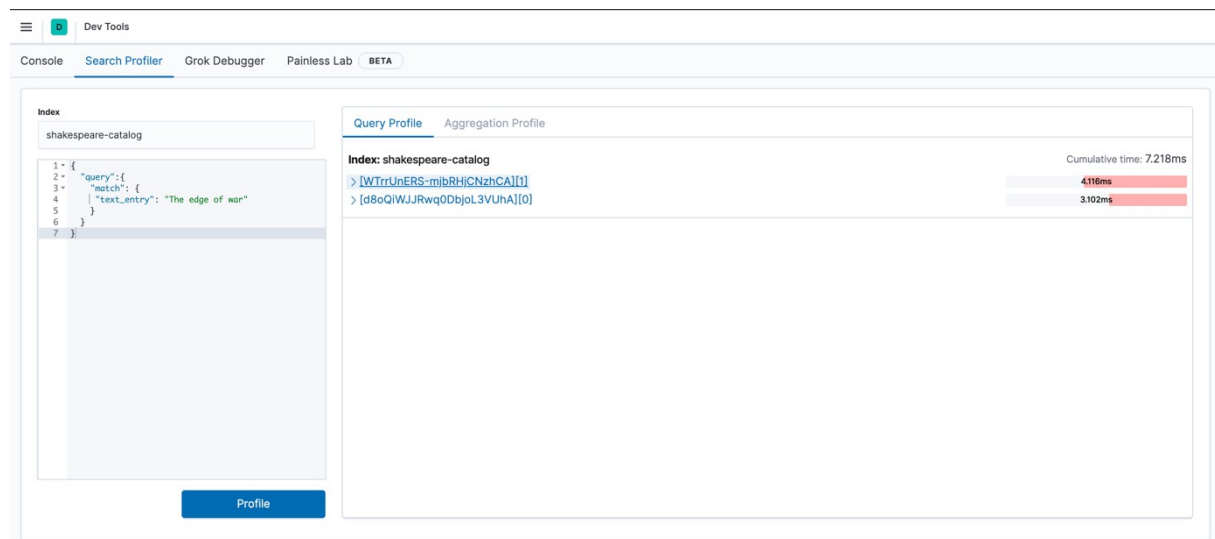
3.3 Do a screenshot of the result of this command :

```
1   green open .apm-agent-configuration           b1gne6LpT4Cv1gKU1YqH4Q 1 1      0     0   416b    208b
2   green open logstash-2015.05.20                 OLhZapqUQxei21kf5Wy0Fg 1 1   4750     0  32.9mb  16.4mb
3   green open .monitoring-es-7-2020.11.17         s1bOZO0qRlSbCW6nkWgJDA 1 1  10869  5227  11.9mb   5.9mb
4   green open .kibana_1                           f17nxrzFQSunh6CaN3oVew 1 1     72     5  20.8mb  10.4mb
5   green open shakespeare-catalog                 I9IkeyNPQcGsbyiaZyY0xQ 2 1 111396     0  58.1mb  19.5mb
6   green open bank                                MbMckCtuS16Nrzv1O7gigQ 1 1   1000     0 758.6kb 379.3kb
7   green open .apm-custom-link                    vOknP7eZTqibFZLeObCpJQ 1 1      0     0   416b    208b
8   green open .kibana_task_manager_1              8JrgiQGvS8KtrkCEOb3q0g 1 1      5    79  202kb   76.7kb
9   green open .monitoring-es-7-2020.11.24         wSEAy5MmQ92uUOFtDo9u6g 1 1   4490    76   6.2mb   3.2mb
10  green open shakespeare-catalog-2               swQX4biyRm6nhINYT0Dr6Q 1 1 111396     0  11.5mb   5.7mb
11  green open shakespeare                         pdAMYVATRDSNiTaW0W_NWA 1 1      0     0   416b    208b
12  green open .kibana-event-log-7.10.0-000001     BwdzcvfOTmGaMhfQv6WtIw 1 1      5     0  54.3kb  27.1kb
13  green open .monitoring-kibana-7-2020.11.24     oy1_Lk8-Thi_GFqddR1yJA 1 1    300     0 469.4kb 293.7kb
14  green open .async-search                       _beEXq7XQHWJANOXk32paQ 1 1      0     0   6.7kb   3.3kb
15  green open logstash-2015.05.18                 ckszjGF0Si-UyUgbOLFTUA 1 1   4631     0  25.2mb  12.6mb
16  green open logstash-2015.05.19                 iOwmw_5LRNaZj1K-KzTXLg 1 1   4624     0  25.5mb  12.7mb
17  green open .monitoring-kibana-7-2020.11.17     M_6AMGnKT6WFxyEfcVSPpQ 1 1    814     0  642kb  304.2kb
18
```

3.4 Do a screenshot of this search query results and add it to your report :

```
  1 ▾ {
  2     "took" : 2,
  3     "timed_out" : false,
  4 ▾   "_shards" : {
  5       "total" : 1,
  6       "successful" : 1,
  7       "skipped" : 0,
  8       "failed" : 0
  9 ▴   },
 10 ▾   "hits" : {
 11 ▾     "total" : {
 12         "value" : 422,
 13         "relation" : "eq"
 14 ▴     },
 15       "max_score" : 12.877999,
 16 ▾     "hits" : [
 17 ▾       {
 18           "_index" : "shakespeare-catalog-2",
 19           "_type" : "_doc",
 20           "_id" : "19",
 21           "_score" : 12.877999,
 22 ▾         "_source" : {
 23             "type" : "line",
 24             "line_id" : 20,
 25             "play_name" : "Henry IV",
 26             "speech_number" : 1,
 27             "line_number" : "1.1.17",
 28             "speaker" : "KING HENRY IV",
 29             "text_entry" : "The edge of war, like an ill-sheathed knife,"
 30 ▴         }
 31 ▴       },
 32 ▾       {
 33           "_index" : "shakespeare-catalog-2",
 34           "_type" : "_doc",
```

```
 32 ▾       {
 33           "_index" : "shakespeare-catalog-2",
 34           "_type" : "_doc",
 35           "_id" : "14215",
 36           "_score" : 9.585962,
 37 ▾         "_source" : {
 38             "type" : "line",
 39             "line_id" : 14216,
 40             "play_name" : "Alls well that ends well",
 41             "speech_number" : 2,
 42             "line_number" : "3.3.7",
 43             "speaker" : "BERTRAM",
 44             "text_entry" : "To the extreme edge of hazard."
 45 ▴         }
 46 ▴       },
 47 ▾       {
 48           "_index" : "shakespeare-catalog-2",
 49           "_type" : "_doc",
 50           "_id" : "55056",
 51           "_score" : 9.585962,
 52 ▾         "_source" : {
 53             "type" : "line",
 54             "line_id" : 55057,
 55             "play_name" : "Loves Labours Lost",
 56             "speech_number" : 130,
 57             "line_number" : "5.2.274",
 58             "speaker" : "BOYET",
 59             "text_entry" : "As is the razors edge invisible,"
 60 ▴         }
```

3.5 Do copy / paste all the queries you wrote in Dev Tools during this lab and add it to your report :

```
PUT /shakespeare/_settings
{
 "settings": {
 "index.blocks.write": true
 }
}

POST shakespeare/_split/shakespeare-catalog
{
 "settings": {
 "index.number_of_shards": 2
 }
}

DELETE shakespeare

GET /shakespeare-catalog/_search
{
 "query": {
 "match": {
 "text_entry": "The edge of war"
 }
 }
}

GET /shakespeare-catalog/_search
{
 "track_total_hits": true,
 "query": {
 "match": {
 "text_entry": "The edge of war"
 }
 }
}

GET /shakespeare-catalog/_search
{
 "query": {
 "match": {
 "text_entry": {
 "query": "The edge of war",
 "operator": "AND"
 }
 }
 }
}
```

```
GET /shakespeare-catalog/_search
{
 "query": {
 "match": {
 "text_entry": {
 "query": "The edge of war",
 "minimum_should_match": "2"
 }
 }
 }
}
GET /shakespeare-catalog/_search
{
 "query": {
 "match": {
 "text_entry": {
 "query": "The edge of war",
 "minimum_should_match": "2"
 }
 }
 },
 "sort": [
 {
 "_score": {
 "order": "asc"
 }
 }
 ]
}
GET /shakespeare-catalog/_mapping

PUT /shakespeare-catalog-2
{
 "mappings": {
 "properties": {
 "speaker": {
 "type": "keyword"
 },
 "play_name": {
 "type": "keyword"
 },
 "line_id": {
 "type": "integer"
 },
 "speech_number": {
 "type": "integer"
 },
 "text_entry": {
 "type": "text",
```

```
  "fields": {
  "english": {
  "type": "text",
  "analyzer": "english"
  }
  }
  }
  }
  }
}

POST _reindex
{
 "source": {
 "index": "shakespeare-catalog"
 },
 "dest": {
 "index": "shakespeare-catalog-2",
 "version_type": "internal"
 }
}

GET /_cat/indices/

GET /shakespeare-catalog-2/_search
{
 "query": {
 "match": {
 "text_entry.english": {
 "query": "The edge of war",
 "minimum_should_match": "2"
 }
 }
 },
 "sort": [
 {
 "_score": {
 "order": "asc"
 }
 }
 ]
}

GET /_analyze/
{
 "analyzer": "default",
 "text":"The edge of war"
}

GET /_analyze/
```

```
{
 "analyzer": "english",
 "text":"The edge of war"
}

GET /shakespeare-catalog-2/_search
{
 "query": {
 "match": {
 "text_entry.english": {
 "query": "The edge of war"
 }
 }
 },
 "sort": [
 {
 "_score": {
 "order": "asc"
 }
 }
 ]
}

GET /shakespeare-catalog-2/_search
{
 "query": {
 "match": {
 "text_entry.english": {
 "query": "The edge of war"
 }
 }
 },
 "size": 20
}
```

3.6 Do a screenshot of your aggregation query results and add it to your report :

```
 1▾ {
 2    "took" : 19,
 3    "timed_out" : false,
 4▾   "_shards" : {
 5    |   "total" : 1,
 6    |   "successful" : 1,
 7    |   "skipped" : 0,
 8    |   "failed" : 0
 9▴   },
10▾   "hits" : {
11▾   |   "total" : {
12    |   |   "value" : 10000,
13    |   |   "relation" : "gte"
14▴   |   },
15    |   "max_score" : null,
16    |   "hits" : [ ]
17▴   },
18▾   "aggregations" : {
19▾   |   "play_names" : {
20    |   |   "doc_count_error_upper_bound" : 0,
21    |   |   "sum_other_doc_count" : 73252,
22▾   |   |   "buckets" : [
23▾   |   |   |   {
24    |   |   |   |   "key" : "Hamlet",
25    |   |   |   |   "doc_count" : 4244
26▴   |   |   |   },
27▾   |   |   |   {
28    |   |   |   |   "key" : "Coriolanus",
29    |   |   |   |   "doc_count" : 3992
30▴   |   |   |   },
31▾   |   |   |   {
32    |   |   |   |   "key" : "Cymbeline",
33    |   |   |   |   "doc_count" : 3958
34▴   |   |   |   },
```

3.7 Do a screenshot of your search & aggregation query results and add it to your report :

```
},
"hits" : {
    "total" : {
        "value" : 10000,
        "relation" : "gte"
    },
    "max_score" : 16.202208,
    "hits" : [
        {
            "_index" : "shakespeare-catalog-2",
            "_type" : "_doc",
            "_id" : "19",
            "_score" : 16.202208,
            "_source" : {
                "type" : "line",
                "line_id" : 20,
                "play_name" : "Henry IV",
                "speech_number" : 1,
                "line_number" : "1.1.17",
                "speaker" : "KING HENRY IV",
                "text_entry" : "The edge of war, like an ill-sheathed knife,"
            }
        },
        {
            "_index" : "shakespeare-catalog-2",
            "_type" : "_doc",
            "_id" : "14215",
            "_score" : 12.496816,
            "_source" : {
                "type" : "line",
                "line_id" : 14216,
                "play_name" : "Alls well that ends well",
                "speech_number" : 2,
                "line_number" : "3.3.7",
                "speaker" : "BERTRAM",
                "text_entry" : "To the extreme edge of hazard."
            }
        },
        {
            "_index" : "shakespeare-catalog-2",
            "_type" : "_doc"
```

```
  },
  "aggregations" : {
    "speakers" : {
      "doc_count_error_upper_bound" : 0,
      "sum_other_doc_count" : 28837,
      "buckets" : [
        {
          "key" : "GLOUCESTER",
          "doc_count" : 596
        },
        {
          "key" : "HAMLET",
          "doc_count" : 551
        },
        {
          "key" : "KING HENRY V",
          "doc_count" : 412
        },
        {
          "key" : "FALSTAFF",
          "doc_count" : 400
        },
        {
          "key" : "IAGO",
          "doc_count" : 294
        },
        {
          "key" : "Clown",
          "doc_count" : 284
        },
        {
          "key" : "MACBETH",
          "doc_count" : 284
        },
        {
          "key" : "KING HENRY VI",
          "doc_count" : 282
        },
        {
          "key" : "BRUTUS",
          "doc_count" : 279
        },
        {
          "key" : "KING RICHARD II",
```

3.8 Do a screenshot of the response in the inspect panel and add it to your report :
Interval: Auto

```
  "took": 2,
  "timed_out": false,
  "_shards": {
    "total": 3,
    "successful": 3,
    "skipped": 0,
    "failed": 0
  },
  "hits": {
    "total": 13992,
    "max_score": null,
    "hits": []
  },
  "aggregations": {
    "2": {
      "buckets": [
        {
          "key_as_string": "2015-05-18T00:00:00.000+02:00",
          "key": 1431900000000,
          "doc_count": 4619
        },
        {
          "key_as_string": "2015-05-19T00:00:00.000+02:00",
          "key": 1431986400000,
          "doc_count": 4622
        },
        {
          "key_as_string": "2015-05-20T00:00:00.000+02:00",
          "key": 1432072800000,
          "doc_count": 4751
        },
```

Interval daily:

```
{
  "took": 2,
  "timed_out": false,
  "_shards": {
    "total": 3,
    "successful": 3,
    "skipped": 0,
    "failed": 0
  },
  "hits": {
    "total": 13992,
    "max_score": null,
    "hits": []
  },
  "aggregations": {
    "2": {
      "buckets": [
        {
          "key_as_string": "2015-05-18T02:00:00.000+02:00",
          "key": 1431907200000,
          "doc_count": 5
        },
        {
          "key_as_string": "2015-05-18T03:00:00.000+02:00",
          "key": 1431910800000,
          "doc_count": 3
        },
        {
          "key_as_string": "2015-05-18T04:00:00.000+02:00",
          "key": 1431914400000,
          "doc_count": 23
        },
        {
          "key_as_string": "2015-05-18T05:00:00.000+02:00",
          "key": 1431918000000,
```
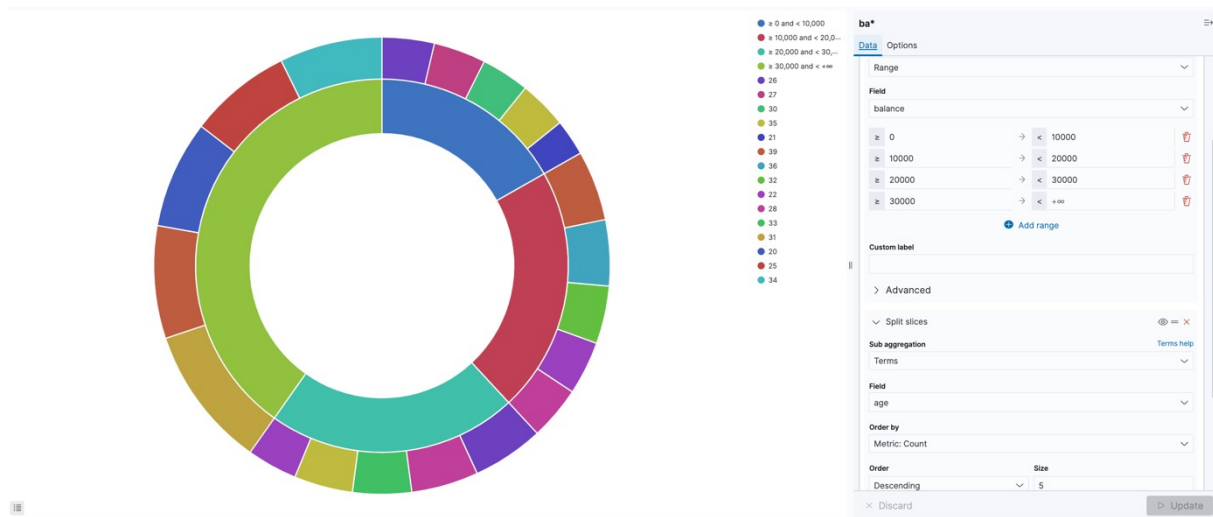
3.9 Do a screenshot of the response in the inspect panel and add it to your report :

```
{
  "took": 3,
  "timed_out": false,
  "_shards": {
    "total": 3,
    "successful": 3,
    "skipped": 0,
    "failed": 0
  },
  "hits": {
    "total": 13992,
    "max_score": null,
    "hits": []
  },
  "aggregations": {
    "2": {
      "buckets": [
        {
          "3": {
            "doc_count_error_upper_bound": 0,
            "sum_other_doc_count": 0,
            "buckets": [
              {
                "key": "ios",
                "doc_count": 2
              },
              {
                "key": "win 8",
                "doc_count": 2
              },
              {
                "key": "win 7",
                "doc_count": 1
              }
            ]
```

3.10 Do a screenshot of the response in the inspect panel and add it to your report :

```
{
  "took": 1,
  "timed_out": false,
  "_shards": {
    "total": 1,
    "successful": 1,
    "skipped": 0,
    "failed": 0
  },
  "hits": {
    "total": 1000,
    "max_score": null,
    "hits": []
  },
  "aggregations": {
    "2": {
      "buckets": {
        "0.0-10000.0": {
          "from": 0,
          "to": 10000,
          "doc_count": 168
        },
        "10000.0-20000.0": {
          "from": 10000,
          "to": 20000,
          "doc_count": 213
        },
        "20000.0-30000.0": {
          "from": 20000,
          "to": 30000,
          "doc_count": 217
        },
        "30000.0-*": {
          "from": 30000,
          "doc_count": 402
        }
      }
    }
  }
}
```

3.11 Do a screenshot of the visualization and settings and add it to your report :

3.12 Do a screenshot of the dashboard and add it to your report